

RATIONAL DESIGN OF p53, AN INTRINSICALLY UNSTRUCTURED PROTEIN, FOR THE FABRICATION OF NOVEL MOLECULAR SENSORS

Melissa L. Geddie, Taryn L. O'Loughlin, Kristen K. Woods¹ and Ichiro Matsumura*

Department of Biochemistry, Center for Fundamental and Applied Molecular Evolution, Emory University School of Medicine, Atlanta, GA

Running Title: p53 as a molecular sensor

*Address Correspondence: Ichiro Matsumura, Department of Biochemistry, Center for Fundamental and Applied Molecular Evolution, Emory University School of Medicine, Rollins Research Center, Room 4119, 1510 Clifton Road, Atlanta, GA 30322, tele: (404) 727-5625, FAX: (404) 727-3452 e-mail address: imatsum@emory.edu

¹Current Address: United States Department of Agriculture, Plant Polymer Research, NCAUR, 1815 N. University Street Peoria, Il, 61604 (Phone: 309 681-6335, Fax: 309 681-6691)

The dominant paradigm of protein engineering is structure-based site-directed mutagenesis. This rational approach is generally more effective for the engineering of local properties, such as substrate specificity, than global ones such as allostery. Previous workers have modified normally unregulated reporter enzymes, including beta-galactosidase, alkaline phosphatase and beta-lactamase, so that the engineered versions are activated (up to 4-fold) by monoclonal antibodies. A reporter that could easily be "re-programmed" for the facile detection of novel effectors (binding or modifying activities) would be useful in high throughput screens for directed evolution or drug discovery. Here we describe a straightforward and general solution to this potentially difficult design problem. The transcription factor p53 is normally regulated by a variety of post-translational modifications. The insertion of peptides into intrinsically unstructured domains of p53 generated variants that were activated up to 100-fold by novel effectors (proteases or antibodies). An engineered p53 was incorporated into an existing high throughput screen for the detection of HIV protease, an arbitrarily chosen novel effector. These results suggest

that the molecular recognition properties of intrinsically unstructured proteins are relatively easy to engineer, and that the absence of crystal structures should not deter the rational engineering of this class of proteins.

INTRODUCTION

Cells generally employ sensor proteins (also called "biosensors" or "switches") to detect chemical stimuli and activate downstream components of signal transduction systems. We seek to fabricate artificial molecular sensors by engineering proteins that are specifically activated when bound or modified by novel effectors. Such sensors have practical utility in high throughput screens for drug discovery or directed protein evolution. They have also proven to be useful as research reagents. For example, two-hybrid systems (1,2) and protein fragment complementation assays (3) couple the interactions of fusion proteins within transgenic cells to the production of signals. Cleverly designed sensors based upon fluorescent

resonance energy transfer (FRET)¹ between Green Fluorescent Protein (GFP) analogues have also enabled the observation of intracellular protein modification events (4), including protein phosphorylation (5) and proteolysis (6). We expect that the utility of engineered proteins sensors will continue to increase as they are deployed as diagnostic reagents (7) and pathogen-activated biotherapeutics (8,9).

The dominant model of protein engineering is structure-based site-directed mutagenesis (10). Reporter proteins are usually selected as starting points for sensor design because their structures have been solved, and because their activities are amenable to high throughput screening. Previous workers have inserted peptide epitopes into beta-galactosidase (11), alkaline phosphatase (12), or beta-lactamase (13). This approach has generally produced catalytically compromised enzymes that are activated up to 4-fold by antibody binding (14), presumably through allosteric mechanisms (7). It nevertheless remains difficult to predict whether the insertion of any peptide epitope into a particular position of a protein will generate the desired antibody-dependent activity. In contrast, natural selection has no bias in favor of proteins that crystallize readily or those with spectroscopically detectable activities. It has in effect generated vast numbers of proteins that are regulated through modification or binding, presumably through parsimonious evolutionary pathways. We therefore considered nature's solutions to the problem of sensor design before formulating our own strategies.

By choosing globular and normally unregulated reporter enzymes such as beta-galactosidase and alkaline phosphatase, protein engineers may be undertaking unnecessarily difficult design problems. In contrast, natural

proteins that participate in signal transduction and gene expression tend to be intrinsically unstructured (15). The unbound forms of these proteins have been described as "beads on a flexible string", where the beads are domains (often molecular recognition elements) connected by linkers (16). We hypothesize that intrinsically unstructured proteins are easy to engineer due to their inherent modularity and relative absence of functional constraint.

Our strategy is therefore to fabricate novel sensors by engineering an intrinsically unstructured protein. Sensors that are activated are preferable to those that are inactivated, since the latter are more likely to produce "false positives" during high-throughput screens. Lim and co-workers previously "re-programmed" the effector-dependence and gating behavior of the neuronal Wiskott-Aldrich syndrome protein (N-WASP) (17), which also contains unstructured regions (according to the DisProt database (18)). N-WASP is modular in design and easily "re-programmed", but its activity (actin polymerization) is not particularly convenient to assay.

We chose the transcription factor p53 as a starting point for several reasons. It is an important tumor suppressor, so its structure and function are well understood (Figure 1). Regions within the N- and C-termini of p53 are thought to be intrinsically unstructured (19,20) and are therefore likely to accommodate almost any insertion. The wild-type p53 remains inactive *in vitro* until the C-terminal 30 amino acids are bound by an antibody, deleted or phosphorylated (21). The exact mechanism of activation remains unclear (22), but can nevertheless be exploited. The sequence-specific DNA-binding activity of p53 can be detected *in vitro* using labeled DNA probes, or *in vivo* using reporter genes within yeast or mammalian cells (Figure 1b).

p53 binds double-stranded DNA with two adjacent copies of the consensus sequence: 5'-PuPuPuC(A/T)(T/A)GPyPyPy-3' (23), and its transcription activation activity in cultured cells is dependent on DNA binding (24). We and others employ the artificial p53CON target sequence, 5'-GGACATGCCCCGGGCATGTCC-3', which was

¹ Abbreviations used: FRET, fluorescent resonance energy transfer; GFP, green fluorescent protein; N-WASP, neuronal Wiskott-Aldrich syndrome protein; p53CON, artificial p53 target sequence (25); HIV, Human Immunodeficiency Virus; IMAC, immobilized metal affinity chromatography; EMSA, electrophoretic mobility shift assay; IPTG, isopropyl-thio-galactopyranoside

isolated from a pool of random sequence double-stranded oligonucleotides (25). Active p53 protein can be expressed in *E. coli*, and is thus inexpensive and easy to produce. The high affinity of p53 for p53CON ($K_D = 5 \times 10^{-10}$ M, (26)) enables highly sensitive assays. Active and latent p53 proteins differ greatly in sequence-specific DNA binding activity. In fact, the "activation factor" we observed of p53 (as high as ~100-fold increase in signal) greatly exceeds those of textbook allosteric enzymes (<70% increase, (27)). The signal produced by p53 *in vitro* remains constant once it reaches equilibrium, so assays are less time-dependent and labor-intensive than those that employ enzymes.

Here we demonstrate the versatility of p53-based molecular sensors. Site-directed insertion mutagenesis and heterologous expression were used to fabricate p53 variants that display peptides recognized by proteases (HIV protease, *Bacillus anthracis* Lethal Factor) or monoclonal antibodies (three different epitopes). All of the p53 variants were specifically activated by their designated effectors in *in vitro* assays. These sensors have immediate utility in high throughput screens, and could potentially be used in other applications (Discussion section). More importantly, this work demonstrates a simple but effective alternative to structure-based site-directed mutagenesis for the fabrication of artificial sensors.

EXPERIMENTAL PROCEDURES

Materials

Expression vectors pET20b+, pET28a+ and pCDF Duet were from Novagen (Madison, WI). The p1+IQ HIV protease expression vector (ATCC #68352) and the human p53 cDNA (ATCC #57254) were obtained from the American Type Culture Collection. The Lethal Factor expression vector, pLF, was a gift from Dr. Stephen Leppla. *E. coli* strain BL21(DE3) Gold/pLysS was from Stratagene (La Jolla, CA). *E. coli* strain Inv \square F' was from Invitrogen (Carlsbad, CA). Gamma labeled P-32 ATP was from MP Biomedicals (Irvine, CA) and Butterfly nitrocellulose membranes from Schleicher and Schuell (Keene, NH). Oligonucleotides were synthesized by IDT (Coralville, IA); the IRD-700

labeled oligo was from LiCor (Lincoln, NE). The anti-p53 p53 monoclonal antibody (pAb1801) and the anti-Lethal Factor antibody (BAL0105) were from AbCam (Cambridge, UK). The HA antibody was from Covance (Princeton, NJ); the HSV antibody and purified Lethal Factor protease were from EMD Biosciences (San Diego, CA). The BigDye 3.1 DNA sequencing and GeneAmp XL long PCR kits were from Applied Biosystems (Foster City, CA).

Construction of p53 variants:

The human *p53 Δ 30* gene from vector p53 Δ 30-pET20b+ (28) was fused to sequence encoding an N-terminal six histidine tag by subcloning *p53 Δ 30* into pET28a+ (Novagen, Madison, WI) using restriction enzymes Nde I and Hind III; the 6his-*p53 Δ 30* was then subcloned back into pET20b+ using Xba I and Hind III. The TEM-1 beta-lactamase gene of the resulting 6his-p53 Δ 30-pET20b+ was replaced with the kanamycin phosphotransferase gene from pET28a+ using Bsp HI, thereby creating 6his-p53 Δ 30-pET20b+(kanR). The 3' end of the full-length wild-type p53 from pHp53B was subcloned into 6his-p53 Δ 30-pET20b+ using Nco I and Bam HI, thereby creating the wild-type 6his-wild-type p53-pET20b+ (kanR) vector.

The p53 to detect HIV protease cleavage (p53/p6) was created by replacing amino acids 360-369 with the p6 site by whole plasmid PCR (29) using the following 5' phosphorylated primers (p6 encoding sequence underlined): 5'-CCTCAGATCACTCTGAAGTCCAAAAAGGG TCAGTCTACC - 3' (p53 p6-1) and 5'-GAAGTTAAAGCTTACTGGCTCCTTCCCAGC CTGGGCATC - 3' (p53 p6-2). The p53/LF10 expression vector was similarly created using site-directed insertion of the recognition sequence at p53 codon 364. The primers used were 5' - CCGTATCCGATGGAAGCTCACTCC AGCCACCTGAAGTCCAA - 3' (LF p53) and 5' - CTCGAGATACACTTTCTTCTGCTCCCCC TGGCTCCTTCC - 3' (LF p53rev).

The p53- Δ 68 mutant was created by deleting the sequence encoding the tetramerization and activation domains of p53 in a whole plasmid PCR (29), using primers

5'-TAGGACGTCGAAGCCGAATTCCAGCACAC TGGCG-3' (p53Δ68) and 5'-ATCCAGTGGTTTCTTCTTTGGCTGGGG-3' (p53Δ68rev). All antibody epitopes were inserted in place of the native pAb1801 epitope (codons 46-55) by whole plasmid PCR, as described above. The constructs were created using the following primers (epitope is underlined): 6his-p53-HSV-Δ68 5' - CCGGAAGATCCGGAAGATGAAGACCCAGG TCCAGATGAAGCTCC -3' (HSVp53out) and 5'-CGCCAGTTCCGGCTGCATCAAATCATCCAT TGCTTGGG ACGG -3' (HSVp53rev); 6his-p53-LF(ab)-Δ68 5' - GCGTTCCGCATGATCGGTGCT CAAATCATCCATTGCTTGGGACGGCAA -3' LFP53rev(AB) and CTGAAAGTGC AGAAAAACGCGCCAGGTCCAGATGAAGCT CCCAGAATG (LFP53out(ab)).

Construction of protease expression vectors

The inducible HIV protease expression vector, p1+IQ (*lacZ*) was constructed as follows. The expression vector p1+IQ (30) was obtained from the American Type Culture Collection. The *lacZ* gene was excised from this vector using Bam HI; the remaining DNA was purified, self-ligated and used to transform *E. coli* strain InvαF'. The P_{BAD}-HIV PR-pCDF expression vectors were constructed in two stages. First, we made the P_{BAD}-pCDF expression vector by subcloning the *araC* repressor and P_{BAD} promoter from pBAD myc his A into pSL1180 using Sph I and Nco I; the P_{BAD} promoter was then subcloned from P_{BAD}-pSL1180 into pCDF Duet using Nco I and Xba I. Second, the HIV protease gene in p1+IQ was PCR amplified, subcloned into pET28a+ using Nde I and Hind III and sequenced to confirm its wild-type identity. The subcloning fused DNA encoding a hexa-histidine tag to its 5' end; this tag does not affect enzyme activity (31). The inactivating D25N mutation was introduced into 6his-HIV PR-pET28 by whole circle PCR using primers 5'-GAAGCTCTATTAATAACAGGAGCAGATG-3' (HIVPR-D25N-62) and 5'-CTTTAGTTGCCCCCTATCTTTATTGTG-3' (HIVPR-62out). The wild-type and D25N variants of the 6his-HIV PR gene were subcloned

from their respective 6his-HIV PR-pET28 plasmids into P_{BAD}-pCDF using Nco I and Xho I.

The Lethal Factor protease gene was cloned from the pLF7 vector. The signal peptide and an internal Nco I site was removed using a two step cloning (based on Park and Leppla (32)). The first PCR reaction used the primers 5' -AAAAAAACCATGGCGGGCGGTTCATGGTGA TGTAGG- 3' (5' LF NcoI) and 5' -TTGAAGGTCCATGCAGTAATATAGAACGG - 3' (LF 2088rev). The second PCR reactions used primers 5'-CCGTTCTATATTACTGCATGGACCTTCAA - 3' (LF 2126) and 5'-TTTTTGGGCCCGGATCCTTATGAGTTAATA ATGAAC - 3' (3' LF BamHI). The two products were combined in a third PCR reaction in which the entire Lethal Factor gene amplified using the external 5' LF NcoI and 3'LF BamHI primers. The Lethal Factor gene was then cloned into the pCDF Duet vector (Novagen) for the *in vivo* assays. All variants were sequenced using the Applied Biosystems Big Dye protocol at the Center for Fundamental and Applied Molecular Evolution (Emory University).

Protein purification

E. coli BL21 (DE3) cells containing the plasmid pLysS were transformed with constructs that expressed the wild-type or engineered p53 genes fused to N-terminal six histidine tags. The transformants were grown at 37°C to mid-log (Abs₆₀₀ = 0.3), then induced with 0.5 mM IPTG and shaken at 23°C for four hours. The cells were spun down and stored as a pellet at -80°C. The cells were lysed by sonication and the insoluble fraction removed by centrifugation. The protein was purified as described in the pET manual (Novagen) except that the binding, wash and elution buffers were replaced with p53 binding buffer (10 mM Tris HCl pH 7.6, 50 mM NaCl, 10 % glycerol, 0.1% Triton X-100, and 5 mM beta-mercaptoethanol) without beta-mercaptoethanol, plus 450 mM NaCl and 0, 60 and 1000 mM imidazole respectively. The eluted p53 proteins were dialyzed in p53 binding buffer (plus 450 mM NaCl) and stored in p53 binding buffer (plus 45% glycerol and 450 mM NaCl) at -20°C. The total protein concentration was quantified using the

Bradford protein assay (Bio-Rad, Hercules, CA). HIV protease was purified and refolded according to an established protocol (31) and protein concentration also quantified using the Bradford protein assay. We used a commercially available fluorogenic substrate to show that our HIV protease was as active as those described in the literature (33).

The p53/p6 expression vector was co-expressed with HIV protease by induction with 0.5 mM IPTG for 4 hours at 23 degrees, while the p53/LF10 expression vector was co-expressed with Lethal Factor by induction with 0.5 mM IPTG for 18 hours at 18 degrees. The wild type and engineered p53 proteins were expressed, either alone or co-expressed with HIV protease (from p1+IQ (30)) or LF protease (from LF-pCDF), and purified by IMAC as described above. The purified proteins (2.0 micromolar) were incubated with 1.0 picomoles N-terminal infrared labeled oligo p53 CON, 50 ng of pLS1180 and 10 micrograms BSA in a final volume of 20 microliters. After a 30 min incubation on ice, the mixtures were loaded onto 4% non-denaturing acrylamide gels and evaluated by EMSA as described below.

EMSA assays

The EMSAs were performed as described (34) except p53 binding buffer was used. The purified proteins (50 nM p53-Δ30, 10 nM p53-Δ68, 10 nM p53-HA-Δ68, 10 nM p53-LF(ab)-Δ68, 20 nM p53-HSV-Δ68) were incubated with 1.0 picomole 5' IRD-700 labeled oligo p53 CON (5'-ATGGACATGCCCCGGGCATGTCC-3' (25)) and 0.25 - 1.0 micrograms of monoclonal antibodies in a total volume of 20 microliters. After a 30 min incubation on ice, the mixtures were loaded onto 4% acrylamide gels (19:1 acrylamide to bis-acrylamide, 0.33X TBE, 0.1% Triton X-100) and run at 200V for 1 h at 4°C. The gels were scanned using the LiCor Odyssey Infrared Imager; the intensities of the pixels within each band were quantified with the associated Odyssey software (version 1.1). The activation factors were calculated by dividing the intensities of the antibody/engineered p53/p53CON complexes by

the intensity of the engineered p53/p53CON complex.

The *in vitro* protease assays were performed using purified p53 (p53/p6, p53/LF10, p53Δ30 or wild-type) and protease (HIV protease(31) or Lethal Factor) proteins. The purified p53 proteins (2 micromolar) were reacted with the HIV protease (10 micromolar) or LF (2 micromolar) proteases (EMD biosciences) for 48 hours in p53 binding buffer at 4 degrees C. Following the incubation, the p53 activity was determined by EMSA as described above.

High throughput assay

The screen for p53 function is based on a method developed by Singh *et al.* (35). *E. coli* strain BL21(DE3) cells carrying the T7 lysozyme expression vector, pLysS, were transformed with the p53 expression constructs and plated on Luria broth (LB) plates containing 34 μg/mL chloramphenicol and 100 μg/mL kanamycin (LB-kan/chl). After 16 hours of growth at 37°C, the colonies were adsorbed onto a nitrocellulose filter and transferred colony side up to LB-kan/chl plates containing 0.5 mM IPTG to induce expression of p53. The colonies adsorbed to the nitrocellulose filter were induced at 23°C for four hours. The cells remaining on the original plate were re-grown into full colonies by a further 8 hour incubation at 37°C.

The cell membranes of the p53-expressing colonies were disrupted with chloroform gas for 15 minutes, giving the intracellular T7 lysozyme access to the peptidoglycan. The remaining manipulations were carried out in p53 binding buffer (described above) at 23°C. The filters were treated with 2.5 units/mL DNase I (in p53 binding buffer plus 10 mM MgCl₂) for 15 minutes, blocked with 5% non-fat dry milk (in p53 binding buffer plus 40 mM Tris-HCl pH 7.6) for an hour, washed three times in binding buffer for 5 minutes each and probed with 20 nM radiolabeled oligonucleotide containing the p53CON sequence (underlined):
 5'-GTGGACATGCCCCGGGCATGTCC (25) (plus 5 μg/mL denatured Salmon Sperm DNA) for one hour. The fluorescent IRD700-p53CON probe apparently interacts non-specifically with endogenous *E. coli* proteins (data not shown), and

is therefore unsuitable for the colony lift screen. The filters were washed four times more for 7.5 minutes each and the quantity of probe bound to each filter was measured using a BAS-1000 Bio-imaging Analyzer System (Fujifilm Medical Systems USA, Stamford, CT).

RESULTS

Protease activation of p53

In vivo activation

We first engineered p53 variants that were specifically activated by HIV protease or the *Bacillus anthracis* Lethal Factor (a metalloprotease). These effectors were “novel” because p53 does not ordinarily recognize or respond to them. Both proteases catalyze the hydrolysis of peptide substrates, but do not overlap in sequence or conformation specificity (36,37). Site-directed insertion mutagenesis was applied to replace p53 codons 360-369, which encode an unstructured spacer upstream of the C-terminal autoinhibitory domain (Figure 1), with sequence encoding the HIV protease substrate (p6, VSFNFPQITL). Similarly the sequence encoding the Lethal Factor substrate (LF10, KKVYPYPME (37)) was inserted at codon 364. The engineered p53 variants (designated p53/p6 and p53/LF10) and the wild-type p53 gene were separately co-expressed in *E. coli* with either HIV protease, LF or no protease.

The hexa-histidine-tagged p53 proteins were purified by Immobilized Metal Affinity Chromatography (IMAC) and analyzed by SDS-Poly-Acrylamide Gel Electrophoresis (SDS-PAGE). The engineered protein migrated more quickly than the wild type (at the same rate as the p53 Δ 30 control, which lacks its C-terminal domain), but only after co-expression with protease (data not shown). Equimolar quantities of the purified proteins were incubated with a double-stranded p53CON target sequence (25) conjugated to a near-infrared dye (IRD700-p53CON); the sequence specific DNA binding activity of each was measured in an Electrophoretic Mobility Shift Assay (EMSA). The bands at the top of the gel reflect protein-DNA complexes, and their intensities are a

measure of activation. Protease co-expression increased the apparent activity of the engineered p53 variants by approximately 30 fold for p53/p6, and a factor of over 100 fold for p53/LF10. (Figure 2a).

In vitro protease activation

We confirmed that the proteases directly activated the engineered p53 variants through *in vitro* assays using purified proteins. The p53/p6, p53/LF10, and wild-type p53 proteins were separately expressed in *E. coli* and purified by IMAC; HIV protease was also separately expressed, mostly in inclusion bodies, solubilized in urea, purified by IMAC and re-folded (31). The purified p53 proteins were reacted with either purified HIV protease or Lethal Factor (EMD Biosciences); SDS-PAGE analysis confirmed the expected differences in migration after reactions with the proteases. EMSA analysis showed that Lethal Factor restored the activity of the engineered p53/LF10 variant to that of p53 Δ 30, while the HIV protease elicited more modest ~2-fold activation in p53/p6 (Figure 2b). The latter activation factor is apparently worse than that of the comparable *in vivo* reaction, most likely because the pH of *in vitro* assay (7.6) was optimized for p53 rather than HIV protease (38).

High-throughput protease screen

Several high throughput p53 assays have been reported (see Discussion), and here we demonstrate the utility of the p53/p6 variant within a semi-*in vivo* filter-lift screen (28). *E. coli* BL21(DE3) Gold/pLysS cells were separately transformed with p53/p6, p53 Δ 30 and wild-type p53 expression vectors. The transformed colonies were filter-lifted onto plates supplemented with 0.5 mM isopropyl-thio-galactopyranoside (IPTG). Expression of the p53 genes were induced at room temperature for four hours; the colonies were lysed by exposure to chloroform gas and probed with a radiolabeled p53CON oligonucleotide. Phosphorimager analysis showed that the p53 Δ 30 protein exhibited ~2-fold greater activity (comparing photoluminescence/mm²) than the wild-type or p53/p6 proteins (Figure 2c), consistent with published *in vitro* results (34).

Next we co-transformed *E. coli* BL21(DE3) Gold/pLysS with the p53/p6

expression vector and either P_{BAD}-HIV PR-pCDF or P_{BAD}-D25N-HIV PR-pCDF. The latter vectors are identical except that the first produces the wild-type protease when induced with L-arabinose while the other produces the catalytically inactive D25N mutant. The double transformants were propagated on LB-kan/spectinomycin/chl plates, then filter-lifted onto plates containing 0.5 mM IPTG or 0.5 mM IPTG plus 0.2% L-arabinose. Colonies expressing the wild-type HIV PR exhibited ~2-fold greater p53 activity than those expressing the inactive D25N protease variant (Figure 2d). We have already used this screen to detect and isolate p53 mutants with increased activity (28); the precision of the assay is more than sufficient to detect a 2-fold difference in activity. This result demonstrates the compatibility of the engineered p53 variants with existing high throughput screens. Both p53 and HIV protease are largely insoluble when expressed in *E. coli*, so the visualization of both activities within individual colony remnants indicates the acute sensitivity of p53-based sensors.

Antibody activation of p53

We also designed p53 variants that are activated by monoclonal antibodies, rather than proteases. These sensors would have immediate utility in high throughput screens for antibodies and other binding effectors. Our strategy is to engineer monomeric p53 variants that are dimerized by antibodies. The exploitation of this second activation mechanism would further demonstrate the modularity of p53. Engineered p53 monomers, which lack the entire tetramerization domain, exhibit 1000-fold less affinity for their DNA targets than does the tetrameric wild-type protein. Engineered p53 dimers bind their DNA targets with one-sixth the affinity of the wild-type tetramer (39). Induced dimerization of the monomer should therefore be easy to detect in EMSA and other binding assays (Figure 1b).

To create a monomeric form of p53, we used site-directed deletion mutagenesis to remove the tetramerization and auto-inhibitory domains (amino acids 324-393) (40). The resulting p53-Δ68 gene was subcloned into pET28, thus fusing it to DNA encoding an N-terminal hexahistidine-tag. We then replaced the pAb1801 epitope (wild-type

p53 amino acids 46-55) of 6his-p53-Δ68 with sequence encoding the HA epitope (YPYDVPDYA), the HSV epitope (QPELAPEDPED), or an epitope to the Lethal Factor (LF(Ab)-STDHAERLKVQKNA) (Figure 1). All four 6his-p53-Δ68 variants were separately expressed in *E. coli*, and purified by IMAC.

The purified p53-Δ68 proteins were separately reacted with pAb1801 or the anti-epitope monoclonal antibodies and the IRD700-p53CON probe. The antibody-p53-DNA complexes were separated from the free DNA by EMSA (Figure 3a). The p53-Δ68 variant had ~80-fold increase in activity upon the addition of the pAb1801 Ab. p53-HA-Δ68 had approximately a 30-fold increase after adding the HA antibody, p53-HSV-Δ68 had ~100-fold increase, and p53-LF(Ab)-Δ68 had a 3-fold increase in activity (Figure 3b). We were not surprised to see two bands at the top of the gels, as both p53CON and the antibodies are capable of multimeric binding. We assayed our p53 variants at concentrations that were barely detectable in the absence of antibody, and would have obtained even higher activation factors if we had employed the sensor at lower concentrations. The differences in the activation factors might be due to differences in the binding affinity of the antibody to the inserted tag. These values are significantly better than previously reported antibody sensors (14), and should enable assays with broader dynamic range.

DISCUSSION

We have designed p53 variants that are specifically activated by HIV protease, Lethal Factor or monoclonal antibodies specific to the HA, HSV and LF epitopes. These p53-based sensors are versatile with respect to both input molecular recognition and output signals. All of these inserted peptides differ significantly in amino acid sequence and conformation (Figure 1b). HIV protease recognizes bent, hydrophobic peptides (36); Lethal Factor recognizes straight, basic peptides (37). Yet, the p53/p6 and p53/LF10 proteins were specifically recognized, cleaved and activated by HIV protease and Lethal Factor respectively. We therefore believe that p53 could easily be “re-programmed” for the detection of

any protease or antibody. The latter is significant because modification-specific antibodies should enable p53-based HTS for kinases, acetylases and other protein modifying activities

With regard to applications, we showed here that p53 can be detected in EMSA assays and in high throughput colony lift assays. These results underscore the versatility of p53, since it is relatively difficult to control the amount of p53 expression within individual *E. coli* colonies. The wild-type p53 protein functions *in vivo* as a sensor of DNA damage, hypoxia, ribosome biogenesis, rNTP depletion, spindle damage, temperature shock, nitric oxide and oncogene activation; these signals are mediated by an array of upstream regulatory proteins (41). *In vivo* p53 assays based on reporter gene activation have been developed in transgenic mammalian (42) and yeast (43) cells. The co-expression of p53-based sensors and engineered p53-binding proteins should similarly enable *in vivo* screens and selections for a wide variety of effectors.

The HIV protease-activated p53 variant also has potential as a therapy or "intracellular vaccine" for AIDS. Expression of the wild-type p53 protein normally leads to repression of transcription from the HIV-1 LTR (viral promoter; (44)), as well as G1 growth arrest or premature apoptosis. The virus normally overcomes these pleiotropic effects by making a protein, Tat, which represses transcription of p53 (45). Expression of the HIV protease-activated variant in HIV-infected cells would have the following virtues as a gene therapy for AIDS. First, it is unlikely that a naturally occurring protein that requires activation by HIV-1 encoded factors will cause side effects. Second, p53 activity has graduated effects (repression of viral transcription, G1 growth arrest,

apoptosis) that are less drastic than those of other "Trojan horse" therapies (8,9). Third, these effects are mediated by cellular factors that inhibit viral replication, so it is unlikely that HIV-1 could evolve immunity against the engineered p53.

We have also shown that p53 can display a variety of peptide sequences at two different locations. Previous workers have reprogrammed the effector-dependence of N-WASP (17). We are therefore very optimistic about the prospects for the rational design of intrinsically unstructured proteins in general. An estimated 25-40% of all amino acid residues are thought to reside in unstructured domains (46), and many play important regulatory roles (16). We were initially reluctant to attempt rational design in the absence of a crystal structure, but now encourage others to take advantage of the modularity and structural permissiveness of this functionally important class of proteins.

ACKNOWLEDGEMENTS

TLO provided purified HIV protease; KKW constructed the Lethal Factor expression vector and did the *in vivo* Lethal Factor experiments; IM constructed the HIV protease-activated p53 alleles and the HIV protease expression vectors; he also developed the HTS (Figure 2c-d). We thank Dr. Stephen Leppla for providing the Lethal Factor gene, and Dr. Andy Ellington for his ideas. We are grateful to Dr. Justin Gallivan, Mr. Shawn Desai, Dr. Marc Ostermeier and Dr. Wayne Patrick for reading the manuscript. This project was supported by NIH/NIAID (1 R21AI054602-01).

REFERENCES

1. Baker, K., Bleczynski, C., Lin, H., Salazar-Jimenez, G., Sengupta, D., Krane, S., and Cornish, V. W. (2002) *Proc Natl Acad Sci U S A* **99**, 16537-16542
2. Drees, B. L. (1999) *Curr Opin Chem Biol* **3**, 64-70.
3. Michnick, S. W. (2003) *Curr Opin Biotechnol* **14**, 610-617
4. Tsien, R. Y. (1998) *Annu Rev Biochem* **67**, 509-544
5. Violin, J. D., Zhang, J., Tsien, R. Y., and Newton, A. C. (2003) *J Cell Biol* **161**, 899-909

6. Heim, R., and Tsien, R. Y. (1996) *Curr Biol* **6**, 178-182
7. Villaverde, A. (2003) *FEBS Lett* **554**, 169-172
8. Liu, S., Redeye, V., Kuremsky, J. G., Kuhnen, M., Molinolo, A., Bugge, T. H., and Leppla, S. H. (2005) *Nat Biotechnol* **23**, 725-730
9. Vocero-Akbani, A. M., Heyden, N. V., Lissy, N. A., Ratner, L., and Dowdy, S. F. (1999) *Nat Med* **5**, 29-33
10. Brannigan, J. A., and Wilkinson, A. J. (2002) *Nat Rev Mol Cell Biol* **3**, 964-970
11. Feliu, J. X., and Villaverde, A. (1998) *FEBS Lett* **434**, 23-27.
12. Brennan, C. A., Christianson, K., La Fleur, M. A., and Mandecki, W. (1995) *Proc Natl Acad Sci U S A* **92**, 5783-5787
13. Legendre, D., Soumillion, P., and Fastrez, J. (1999) *Nat Biotechnol* **17**, 67-72.
14. Feliu, J. X., Ferrer-Miralles, N., Blanco, E., Cazorla, D., Sobrino, F., and Villaverde, A. (2002) *Biochim Biophys Acta* **1596**, 212-224
15. Tompa, P. (2005) *FEBS Lett* **579**, 3346-3354
16. Dyson, H. J., and Wright, P. E. (2005) *Nat Rev Mol Cell Biol* **6**, 197-208
17. Dueber, J. E., Yeh, B. J., Chak, K., and Lim, W. A. (2003) *Science* **301**, 1904-1908
18. Vucetic, S., Obradovic, Z., Vacic, V., Radivojac, P., Peng, K., Iakoucheva, L. M., Cortese, M. S., Lawson, J. D., Brown, C. J., Sikes, J. G., Newton, C. D., and Dunker, A. K. (2005) *Bioinformatics* **21**, 137-140
19. Bell, S., Klein, C., Muller, L., Hansen, S., and Buchner, J. (2002) *J Mol Biol* **322**, 917-927
20. Lee, H., Mok, K. H., Muhandiram, R., Park, K. H., Suk, J. E., Kim, D. H., Chang, J., Sung, Y. C., Choi, K. Y., and Han, K. H. (2000) *J Biol Chem* **275**, 29426-29432
21. Hupp, T. R., and Lane, D. P. (1994) *Curr Biol* **4**, 865-875
22. Ahn, J., and Prives, C. (2001) *Nat Struct Biol* **8**, 730-732.
23. el-Deiry, W. S., Kern, S. E., Pietenpol, J. A., Kinzler, K. W., and Vogelstein, B. (1992) *Nat Genet* **1**, 45-49
24. Kern, S. E., Pietenpol, J. A., Thiagalingam, S., Seymour, A., Kinzler, K. W., and Vogelstein, B. (1992) *Science* **256**, 827-830
25. Funk, W. D., Pak, D. T., Karas, R. H., Wright, W. E., and Shay, J. W. (1992) *Mol Cell Biol* **12**, 2866-2871
26. Hall, A. R., and Milner, J. (1995) *Oncogene* **10**, 561-567
27. Berg, J. M., Tymoczko, J. L., and Stryer, L. (2002), Fifth Ed., pp. 266, W.H. Freeman and Company, New York
28. Matsumura, I., and Ellington, A. D. (1999) *Protein Sci* **8**, 731-740.
29. Geddie, M. L., and Matsumura, I. (2004) *J Biol Chem* **279**, 26462-26468
30. Baum, E. Z., Bebernitz, G. A., and Gluzman, Y. (1990) *Proc Natl Acad Sci U S A* **87**, 10023-10027.
31. Vickrey, J. F., Logsdon, B. C., Proteasa, G., Palmer, S., Winters, M. A., Merigan, T. C., and Kovari, L. C. (2003) *Protein Expr Purif* **28**, 165-172
32. Park, S., and Leppla, S. H. (2000) *Protein Expr Purif* **18**, 293-302
33. Matayoshi, E. D., Wang, G. T., Krafft, G. A., and Erickson, J. (1990) *Science* **247**, 954-958
34. Hupp, T. R., Meek, D. W., Midgley, C. A., and Lane, D. P. (1992) *Cell* **71**, 875-886.
35. Singh, H., Clerc, R. G., and LeBowitz, J. H. (1989) *Biotechniques* **7**, 252-261

36. Prabu-Jeyabalan, M., Nalivaika, E., and Schiffer, C. A. (2002) *Structure (Camb)* **10**, 369-381
37. Turk, B. E., Wong, T. Y., Schwarzenbacher, R., Jarrell, E. T., Leppla, S. H., Collier, R. J., Liddington, R. C., and Cantley, L. C. (2004) *Nat Struct Mol Biol* **11**, 60-66
38. Louis, J. M., Wondrak, E. M., Kimmel, A. R., Wingfield, P. T., and Nashed, N. T. (1999) *J Biol Chem* **274**, 23437-23442
39. Weinberg, R. L., Veprintsev, D. B., and Fersht, A. R. (2004) *J Mol Biol* **341**, 1145-1159
40. Hansen, S., Hupp, T. R., and Lane, D. P. (1996) *J Biol Chem* **271**, 3917-3924
41. Harris, S. L., and Levine, A. J. (2005) *Oncogene* **24**, 2899-2908
42. Zhang, W. W., Labrecque, S., Azoulay, E., Dudley, R., and Matlashewski, G. (2001) *J Biotechnol* **84**, 79-86
43. Ishioka, C., Frebourg, T., Yan, Y. X., Vidal, M., Friend, S. H., Schmidt, S., and Iggo, R. (1993) *Nat Genet* **5**, 124-129
44. Subler, M. A., Martin, D. W., and Deb, S. (1992) *J Virol* **66**, 4757-4762.
45. Li, C. J., Wang, C., Friedman, D. J., and Pardee, A. B. (1995) *Proc Natl Acad Sci U S A* **92**, 5461-5464.
46. Tompa, P. (2003) *Bioessays* **25**, 847-855

FIGURE LEGENDS

Figure 1: p53 as a universal translator. **(a.)** The domain structure of the wild-type p53 protein is shown with the unstructured regions represented as curvy lines. The peptide sequences inserted into the p53 protein are shown in the table below. **(b.)** The versatility of p53 as a sensor is represented as a flow-chart.

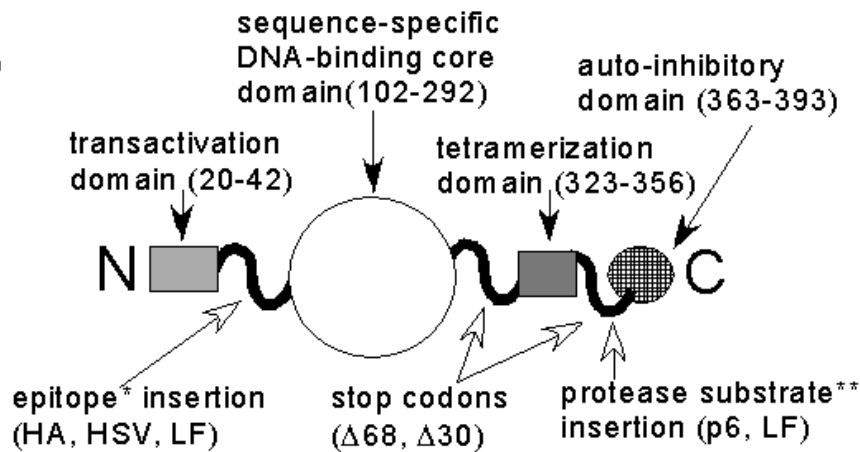
Figure 2: p53 is activated *in vivo* and *in vitro* by HIV and Lethal Factor protease.

(a.) The engineered p53 variant or a wild-type control, was co-expressed in *E. coli* with (or without) HIV or LF protease. The p53 proteins were purified by Immobilized Metal Affinity Chromatography (IMAC), and their sequence-specific DNA binding activity was measured in an Electrophoretic Mobility Shift Assay (EMSA). **(b.)** The engineered and wild-type p53 proteins were separately expressed in *E. coli* and purified by IMAC. Each of the full-length p53 proteins were reacted with purified HIV protease (31) or purified LF protease for 48 hours in p53 binding buffer at 4 degrees C. The DNA binding activity of each variant was determined in an EMSA. **(c.)** The wild-type, activated (p53 Δ 30) and engineered (p53/p6) genes were induced within *E. coli* colonies. The colonies were lysed and probed with a radiolabeled double-stranded oligonucleotide encoding a p53 binding site (p53CON). **(d.)** *E. coli* were co-transformed with a p53/p6 expression vector and either P_{BAD}-wild-type HIV PR-pCDF (bottom row) or P_{BAD}-D25N HIV PR-pCDF (top row). The resulting colonies were induced with 0.5 mM IPTG (p53/p6 only, left column) or 0.5 mM IPTG plus 0.2% L-arabinose (p53/p6 and HIV PR, right column).

Figure 3: Engineered p53 monomers are specifically activated by monoclonal antibodies. **(a.)** The monomeric p53- Δ 68 variants were expressed with N-terminal hexahistidine-tags in *E. coli*. We then replaced the pAb1801 epitope (wild-type p53 codons 46-55) of 6his-p53- Δ 68 with sequence encoding the HA epitope, the HSV epitope, or an epitope to the Lethal Factor from *Bacillus anthracis* (see experimental methods). These p53 variants were purified and reacted with pAb1801 or the anti-HA, HSV or LF monoclonal antibodies and an IRD700-labeled double stranded p53 target sequence (p53CON). The antibody-p53-DNA complexes were separated from the free DNA by an electrophoretic mobility shift assay (EMSA). **(b.)** Quantification of the activation factors (n=3). The pixel intensities of bands

containing antibody/engineered p53/p53CON complexes were divided by those of bands associated with the corresponding engineered p53/p53CON (no antibody) complex.

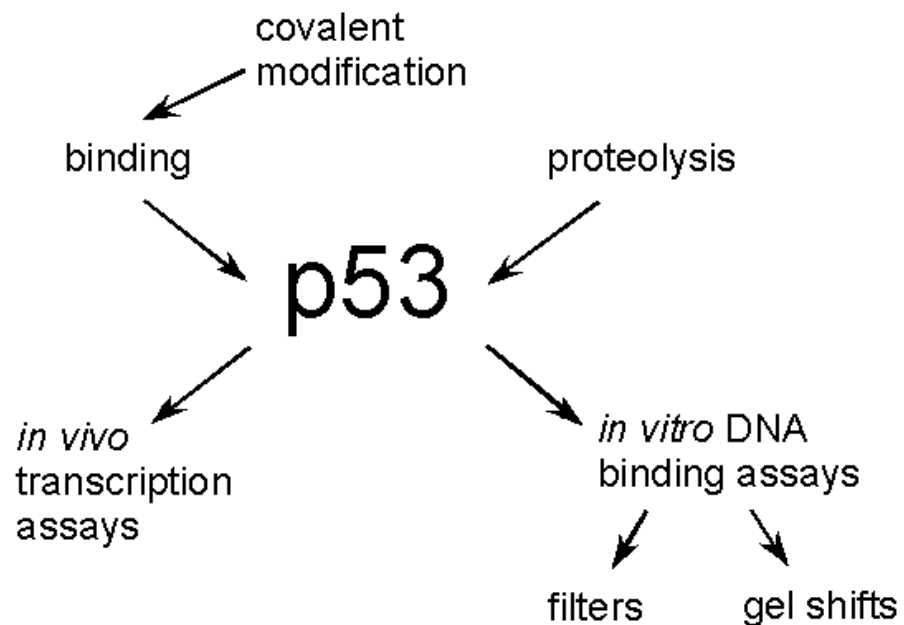
a.



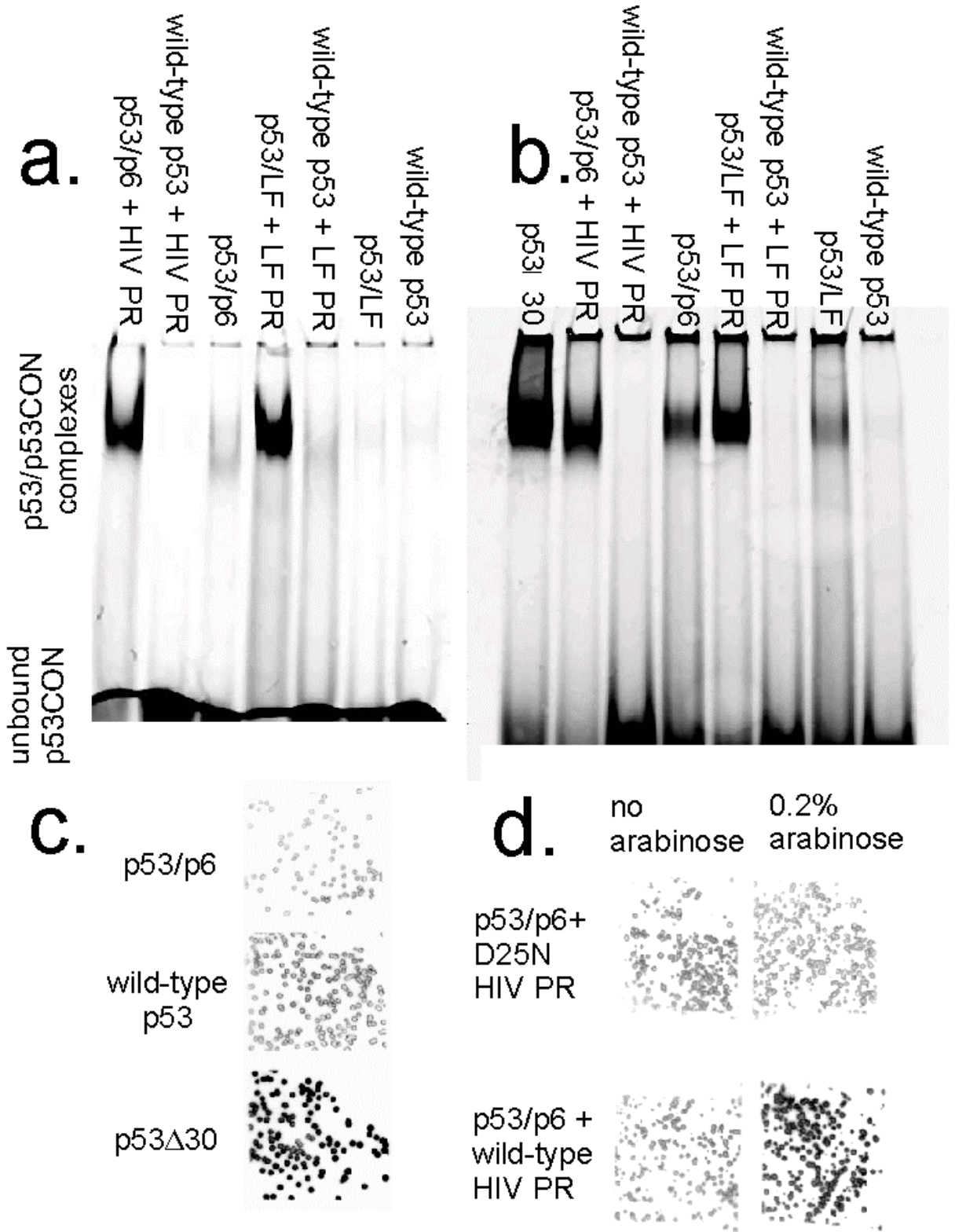
antibody	*epitope	origin
pAB1801	SPDDIENWFT	native p53
anti-HA	YPYDVPDYA	influenza hemagglutinin
anti-HSV	QPELAPEDPED	herpes simplex virus
anti-LF	STDHAERLKVQKNA	<i>B. anthracis</i> Lethal Factor

protease	**substrate peptide	origin
HIV	VSFNFPQITL	p6 of HIV-1 polyprotein
LF	KKVYPYPME	mitogen activate kinase kinase

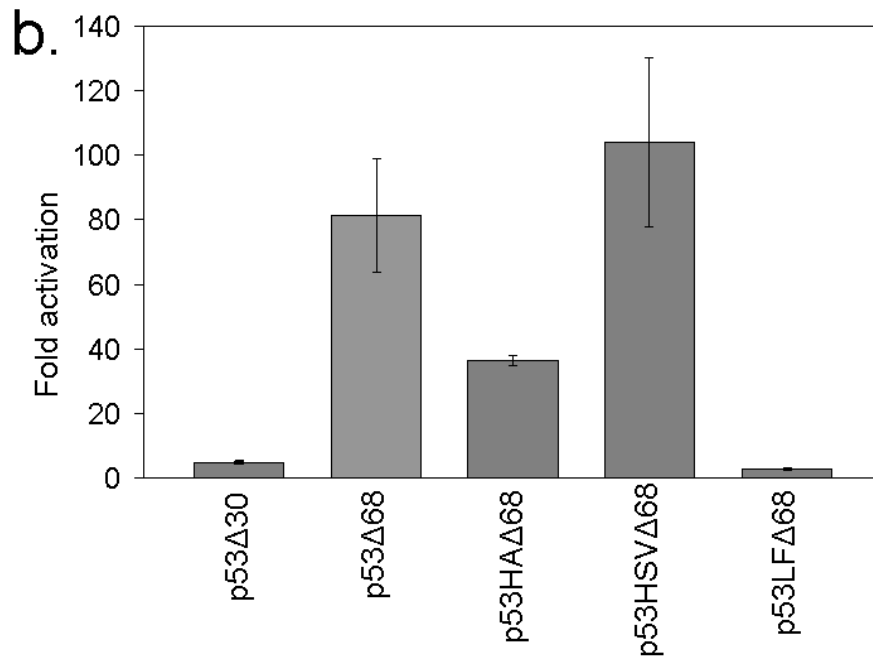
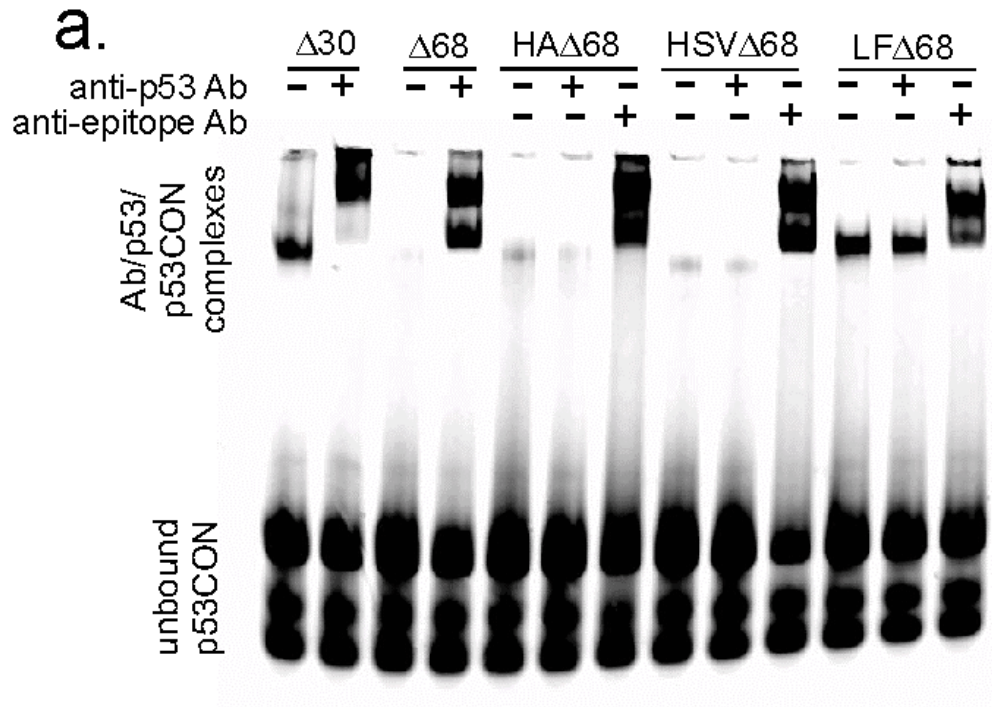
b.



Geddie *et al.*, Fig. 1



Geddie et al., Figure 2



Geddie et al., Figure 3